

DRUG REPURPOSING: TARGETING BCL PROTEIN FAMILY FOR CANCER– AN IN-SILICO APPROACH

Submitted in partial fulfilment for the requirements for the award of

Bachelor of Technology in Biotechnology

An Inter-disciplinary Project

by

JEEVANKUMAAR V (41230023)

PUGAZHESWARAN M (41320041)

SRIRAM A (41230062)



DEPARTMENT OF BIOTECHNOLOGY

SCHOOL OF BIO AND CHEMICAL ENGINEERING

SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY) CATEGORY-1 UNIVERSITY BY UGC

Accredited with grade “A++” by NAAC |12B Status by UGC| Approved by AICTE

JEPPIAAR NAGAR, RAJIV GANDHI SALAI, CHENNAI – 600119

MAY 2024



SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)

CATEGORY-I UNIVERSITY BY UGC

Accredited "A++" Grade by NAAC | 12B Status by UGC | Approved by AICTE

www.sathyabama.ac.in

DEPARTMENT OF BIOTECHNOLOGY

BONAFIDE CERTIFICATE

This is to certify that this Project Report is the bonafide work of **Jeevankumaar V (41230023)**, **Pugazheswaran M (41230041)**, **Sriram A (41230062)** who carried out the Interdisciplinary Project entitled "**DRUG REPURPOSING: TARGETING BCL PROTEIN FAMILY FOR CANCER- AN IN-SILICO APPROACH**" under our supervision from February 2024 to May 2024.

Internal Guide

Dr. B DEVI M.Tech, Ph.D

Head of the Department

Dr. V RAMESH KUMAR M.Phil, Ph.D

Submitted for Viva-voce Examination held on 06.05.2024

Internal Examiner

External Examiner

DECLARATION

We, **JEEVANKUMAAR V** (Register Number: 41320023) **SRIRAM A** (Register Number 41320062) and **PUGAZHESWARAN M** (Register Number 41230041) hereby declare that the Project Report entitled "**DRUG REPURPOSING: TARGETING BCL PROTEIN FAMILY FOR CANCER- AN IN-SILICO APPROACH**" was done by us under the guidance of **Dr. Devi M.Tech, Ph.D.** is submitted in partial fulfilment of the requirements for the award of Bachelor of Technology degree in Biotechnology.



DATE: 06.05.2024

SIGNATURE OF THE CANDIDATE

ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to the Board of Management of **SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY** for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

I convey my sincere thanks to **Dr. V Ramesh Kumar**, Head of the Department, Department of Biotechnology for providing me necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my project Guide **Dr. B Devi**, Assistant Professor, Department of Biotechnology for her valuable guidance, suggestions and constant encouragement paved the way for the successful completion of my project work.

I also express my thanks to **all Teaching and Non-teaching staff members** of the Department of Biotechnology who were helpful in many ways for the completion of the project work.

ABSTRACT

Cancer continues to be a global health challenge, necessitating innovative strategies for therapeutic intervention. Drug repurposing, the process of identifying new uses for existing drugs, offers a promising avenue for expedited drug development. The BCL protein family, comprising both anti-apoptotic and pro-apoptotic members, plays a pivotal role in cancer cell survival and apoptosis evasion. Targeting these proteins presents an attractive therapeutic opportunity.

In this study, we employed an in-silico approach to explore the potential of repurposing existing drugs to target the BCL protein family for cancer treatment. Leveraging computational techniques such as molecular docking, molecular dynamics simulations, and structure-based drug design, we evaluated the binding affinities and interactions of known drugs with various members of the BCL family.

Our findings reveal several candidate drugs with high binding affinities to specific BCL proteins, suggesting their potential as repurposed agents for cancer therapy. Furthermore, we elucidate the structural basis of drug-protein interactions, providing insights into the mechanisms underlying their efficacy.

This in-silico exploration not only identifies promising candidates for further experimental validation but also underscores the utility of computational methods in drug repurposing efforts. By repurposing existing drugs to target the BCL protein family, this approach offers a cost-effective and expedited route to novel cancer therapeutics, with the potential to significantly impact patient outcomes.

LIST OF CONTENTS

Chapter Number	TITLE	Page Number
	Abstract	v
1	INTRODUCTION	1
2	REVIEW OF LITERATURE	3
3	MATERIALS AND METHODS	19
3.1	Bioinformatic Tools Used	19
3.2	Methodology	24
4	RESULTS AND DISCUSSION	28
4.1	Results	28
4.2	Discussion	29
5	SUMMARY AND CONCLUSION	32
6	REFERENCES	33

LIST OF FIGURES

Figure Number	TITLE	Page Number
2.1	Categorization of BCL-2 Family Proteins	6
2.2	2D Structure of Paclitaxel	9
2.3	2D Structure of Navitoclax	10
2.4	2D Structure of TAK-901	12
4.1	Target Protein 3D Structure with Active Sites	28
4.2	Docking Visualisation of Lumateperone to BCL2	29
4.3	Docking Visualisation of Pexidartinib to BCL2	29
4.4	Docking Visualisation of Regorafenib to BCL2	30
4.5	Docking Visualisation of Imidurea to BCL2	30

CHAPTER 1

INTRODUCTION

Cancer remains one of the most formidable challenges to global public health, with its incidence and mortality rates steadily rising worldwide. Despite significant advancements in understanding the molecular mechanisms driving oncogenesis and the development of targeted therapies, the quest for effective cancer treatments continues. Conventional drug discovery and development pipelines are often lengthy, resource-intensive, and fraught with high attrition rates, highlighting the urgent need for innovative approaches to expedite the identification and validation of novel therapeutics. In this context, drug repurposing emerges as a promising strategy to accelerate the availability of new treatments for cancer.

Drug repurposing, also known as drug repositioning or reprofiling, involves the identification of new therapeutic indications for existing drugs that have already undergone safety testing in humans. Unlike *de novo* drug discovery, which typically involves the time-consuming processes of target identification, lead optimization, and preclinical and clinical testing, repurposing leverages the wealth of existing pharmacological knowledge and clinical data to identify candidate drugs with potential efficacy against different diseases, including cancer.

One intriguing avenue for drug repurposing in cancer therapy involves targeting the BCL protein family, a group of key regulators of apoptosis (programmed cell death) and cell survival. The BCL family comprises both pro-apoptotic (e.g., BAX, BAK, BID) and anti-apoptotic (e.g., BCL-2, BCL-XL, MCL-1) members, which play crucial roles in determining cell fate by modulating the intrinsic apoptotic pathway. Dysregulation of BCL proteins is a hallmark of many cancers, contributing to tumor progression, treatment resistance, and poor clinical outcomes. Therefore, targeting these proteins represents an attractive therapeutic strategy for cancer intervention.

In recent years, computational methods have gained prominence in drug discovery and development, offering a cost-effective and time-efficient means to screen large compound libraries, predict drug-target interactions, and optimize lead compounds. In particular, *in-silico* approaches, which utilize computer simulations and modeling techniques to study molecular interactions at the atomic level, have emerged as powerful tools for rational drug design and repurposing.

In this study, we propose an in-silico approach to identify and prioritize candidate drugs for repurposing against the BCL protein family for cancer therapy. By integrating computational techniques such as molecular docking, molecular dynamics simulations, and structure-based drug design, we aim to systematically evaluate the binding affinities and interactions of existing drugs with various members of the BCL family. Through this approach, we seek to elucidate the structural basis of drug-protein interactions and identify lead compounds with the potential for further experimental validation.

By harnessing the power of computational methods and leveraging existing pharmacological knowledge, our study aims to expedite the discovery of novel therapeutics for cancer by repurposing existing drugs to target the BCL protein family. Through this innovative approach, we hope to pave the way for the development of more effective and accessible treatments for cancer patients, ultimately improving clinical outcomes.

CHAPTER 2

REVIEW OF LITERATURE

The Bcl-2 family proteins play a crucial role in cancer by promoting cell survival through the inhibition of apoptosis (Kang & Reynolds, 2009). Targeting these proteins has become essential in cancer therapy due to the evasion of cell death being a hallmark of cancer (Gong, 2023). The anti-apoptotic Bcl-2 family proteins are attractive targets for anticancer drugs as they protect tumor cells from apoptosis under various pressures (Qian *et al.*, 2022). Inhibition of specific members of the anti-apoptotic Bcl-2 family, such as BCL-2, BCL-XL, and BCL-w, enhances apoptosis and facilitates cancer cell clearance (Al-Zabeeby *et al.*, 2018). Inhibitors like Navitoclax and Venetoclax have shown efficacy in promoting cancer cell clearance by enhancing apoptosis (Al-Zabeeby *et al.*, 2018).

By mimicking the action of BH3-only proteins, which are natural antagonists of pro-survival Bcl-2 proteins, drugs like ABT-737 have been developed to overcome chemoresistance in malignant cells (Delft *et al.*, 2006).

In silico approaches have been instrumental in drug discovery and optimization, including the design of interfering peptides against protein-protein interactions (Hashemi *et al.*, 2021). Computational insights have rationalized and predicted crucial drug-target relationships, aiding in drug repurposing for cancer treatment (Cova *et al.*, 2019). Furthermore, in silico models have been used to study the efficacy of therapies, such as personalized cancer treatments, by considering cancer cell morphology and heterogeneity (Hamis *et al.*, 2019). The development of small molecule inhibitors targeting Bcl-2 family proteins has shown promise in pancreatic cancer therapy (Masood *et al.*, 2011). In silico studies have also been utilized to identify potential drug targets and pathways in breast cancer subtypes, emphasizing the importance of pathway cross-talk inhibition (Cava *et al.*, 2018). Additionally, in silico analysis has been employed to study the interactions of miRNAs and polymorphisms in breast cancer patients, providing insights into potential therapeutic targets (Lehmann *et al.*, 2020).

The therapeutic potential of directly inhibiting pro survival proteins was demonstrated with navitoclax, a selective inhibitor of BCL-2 and BCL-X(L), which has shown clinical efficacy in BCL-2-dependent hematological cancers (Souers *et al.*, 2013). Similarly, the anti-apoptotic Bcl-2 family proteins have been identified as therapeutic targets in cancer therapy, with the inhibition of these proteins expected to induce apoptosis in cancer cells (Kang & Reynolds, 2009). Synthetic inhibitors of antiapoptotic Bcl-2 proteins have been a focus of research,

detailing their binding modes, selectivities, and potential for cancer therapy (Yap et al., 2016).

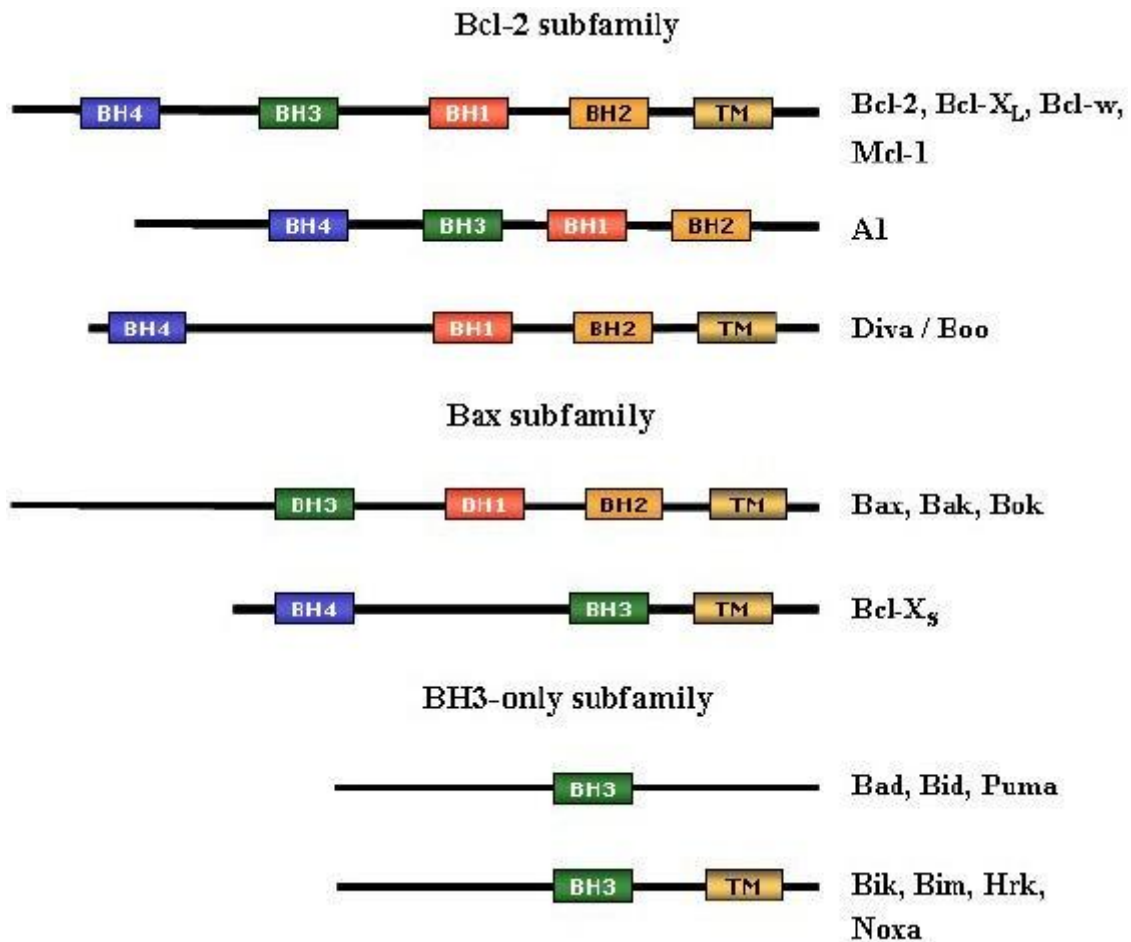
Many studies have been done of the six members of bcl 2 protein for compromising of the both anti-apoptotic and pro-apoptotic members. They all manifest a notably similar folding structure despite an inclusive aberration in amino acid sequence and function. The 3-D structures of Bcl-2 family members consist of two central, principal hydrophobic alpha-helices surrounded by six or seven amphipathic alpha-helices of varying lengths. A long, unstructured loop is present between the first two alpha-helices. The structures of the Bcl-2 proteins show a striking similarity to the overall fold of the pore-forming domains of bacterial toxins. This finding led to experiments which demonstrated that Bcl-x(L), Bcl-2, and Bax all form pores in artificial membranes. A prominent hydrophobic groove is present on the surface of the anti-apoptotic proteins. This groove is the binding site for peptides that mimic the BH3 region of various pro-apoptotic proteins such as Bak and Bad. Structures of Bcl-x(L) in complex with these BH3 peptides showed that they bind as an amphipathic alpha-helix and make extensive hydrophobic contacts with the protein. These data have not only helped to elucidate the interactions important for hetero-dimerization of Bcl-2 family members but have also been used to guide the discovery of small molecules that block Bcl-x(L) and Bcl-2 function. In the recently determined structure of the anti-apoptotic Bcl-w protein, the protein was also found to have a hydrophobic groove on its surface capable of binding BH3-containing proteins and peptides. However, in the native protein an additional carboxy-terminal alpha-helix interacts with the hydrophobic groove.

This is reminiscent of how the carboxy-terminal alpha-helix of the pro-apoptotic protein Bax binds into its hydrophobic groove. This interaction may play a regulatory role and for Bax may explain why it is found predominantly in the cytoplasm prior to activation. The hydrophobic groove of the pro-apoptotic protein, Bid protein, is neither as long nor as deep as that found in Bcl-x(L), Bcl-2, or Bax. In addition, Bid contains an extra alpha-helix, which is located between alpha1 and alpha2 with respect to Bcl-x(L), Bcl-2, and Bax. Although there are still many unanswered questions regarding the exact mechanism by which the Bcl-2 family of proteins modulates apoptosis, structural studies of these proteins have deepened our understanding of apoptosis on the molecular level.

Table 2.1 Members of the Bcl-2 Protein Family

Genes	Role in apoptosis	Function	Localization
Bcl-2 [B-cell Lymphoma - Extra]	Inhibits	Prevents apoptosis by inactivating BAX and BAK [unknown mechanism]	Outer mitochondrial membrane
Bcl-XL [B-cell Lymphoma - Extra]	Inhibits	Prevents Apoptosis by rendering pores non-permeable (thus no pro-apoptotic molecules are released from mitochondria)	Outer mitochondrial membrane
BAX [Bcl-2 Associated X Protein]	Promotes	Opposes Bcl-2 * Forms Oligomeric pore in MOM (Mitochondrial Outer Membrane) * Cytochrome c released * Opening of VDAC (voltage-dependent anion channel)	cytoplasm(inactive), Outer mitochondrial membrane (Active)
BAK [Bcl-2 Homologous Antagonist Killer]	Promotes	* It is tightly bound to MOM even in absence of Apoptotic signal. * Same functions along with BAX	cytoplasm(inactive), outer mitochondrial membrane (active)
BAD [Bcl-2 Associated Death Promoter]	Phosphorylated BAD (Anti - Apoptotic)	No interaction with Bcl-2, prevents Apoptosis	Cytoplasm
BID [BH3 - Interacting Domain Death Agonist]	Pro-Apoptotic	Insertion of BAX into MOM	

Figure 2.1 Categorization of BCL-2 Family Proteins (Roset *et al.*, 2007)



There are still several other pro-apoptotic genes in the Bcl-2 family are BIM (Bcl-2- like protein 11), PUMA (p53 upregulated modulator of apoptosis), NOXA (Phorbol- 12-myristate-13-acetate-induced protein 1).

Based on the domains in the protein, there are three Classes of Bcl-2 Proteins;

- 1) Anti-Apoptotic: Bcl-2 Protein - Bcl-2, Bcl-XL - Has all 4 domains: BH1, BH2, BH3, BH4
- 2) Pro-Apoptotic: BH 123 Protein - BAX, BAK - Has 3 domains: BH1, BH2, BH3
- 3) Pro-Apoptotic: BH3-Only Protein - BAD, BIM, BID, PUMA, NOXA - Has only one domain

The pro-apoptotic members of the BCL2 family are responsible for initiating apoptosis by inducing mitochondrial outer membrane permeabilization (MOMP), leading to the release of cytochrome c and other pro-apoptotic factors from the mitochondria, which ultimately results in caspase activation and cell death. On the other hand, the anti-apoptotic members of the BCL2 family counteract the pro-apoptotic proteins by inhibiting MOMP and preventing cytochrome c release, thereby promoting cell survival. The balance between pro-apoptotic and anti-apoptotic BCL2 family proteins is critical in determining the fate of a cell. Under normal physiological conditions, the anti-apoptotic members keep the pro-apoptotic proteins in check, allowing the cell to survive. However, in response to various cellular stresses, such as DNA damage, oxidative stress, or growth factor withdrawal, the pro-apoptotic proteins are activated, leading to MOMP and initiation of apoptosis. The precise regulation of this interplay is essential for maintaining cellular homeostasis and ensuring proper development, tissue repair, and immune response. The dysregulation of pro-apoptotic and anti-apoptotic BCL2 family proteins is implicated in many diseases, including cancer, neurodegenerative diseases, autoimmune diseases, and cardiovascular diseases. Alterations in the expression, function, or interaction of BCL2 family proteins can shift the balance towards either cell survival or cell death, contributing to disease pathogenesis. Therefore, understanding the complex interplay between pro-apoptotic and anti-apoptotic BCL2 family proteins is crucial for deciphering the molecular mechanisms of apoptosis and developing therapeutic strategies for diseases associated with aberrant apoptosis.

2.2 ROLE OF BCL FAMILY PROTEINS IN CANCER:

- 1) Pro-survival Role: In cancer, anti-apoptotic Bcl-2 family members like Bcl-2 promote cell survival by inhibiting apoptosis, for allowing cancer cells to evade cell death and proliferate uncontrollably.
- 2) Therapeutic Target: Bcl-2 inhibitors have been developed as targeted therapies for various cancers, disrupting the pro-survival function of Bcl-2 proteins and inducing apoptosis in cancer cells. This approach shown promise in treating haematological malignancies and certain solid tumors.

3) BCL-G: A less typical member of the Bcl-2 family, BCL-G, has been studied extensively over the past 20 years. Its unique properties and functions make it an intriguing subject for further research in the context of cancer and other diseases.

2.2 APPROVED DRUGS USED FOR UNDERSTANDING THE MECHANISM

1) PACLITAXEL

2) NAVITOCLAX

3) TAK-901

2.2.1 PACLITAXEL (DB01229):

Paclitaxel interferes with the normal function of microtubule growth. Whereas drugs like colchicine cause the depolymerization of microtubules in vivo, paclitaxel arrests their function by having the opposite effect; it hyper-stabilizes their structure. This destroys the cell's ability to use its cytoskeleton in a flexible manner. Specifically, paclitaxel binds to the β subunit of tubulin.

Tubulin is the "building block" of microtubules, and the binding of paclitaxel locks these building blocks in place. The resulting microtubule/paclitaxel complex does not have the ability to disassemble. This adversely affects cell function because the shortening and lengthening of microtubules (termed dynamic instability) is necessary for their function as a transportation highway for the cell.

Chromosomes, for example, rely upon this property of microtubules during mitosis. Further research has indicated that paclitaxel induces programmed cell death (apoptosis) in cancer cells by binding to an apoptosis stopping protein called Bcl-2 (B-cell leukemia 2) and thus arresting its function.

Chemical Formula of Paclitaxel - $C_{47}H_{51}NO_{14}$

PROPERTY	VALUE	SOURCE
Water Solubility	0.00556 mg/mL	ALOGPS
logP	3.2	ALOGPS
logP	3.54	Chemaxon
logS	-5.2	ALOGPS
pKa (Strongest Acidic)	11.9	Chemaxon
pKa (Strongest Basic)	-1.2	Chemaxon
Physiological Charge	0	Chemaxon
Hydrogen Acceptor Count	10	Chemaxon
Hydrogen Donor Count	4	Chemaxon
Polar Surface Area	221.29 Å ²	Chemaxon
Rotatable Bond Count	14	Chemaxon
Refractivity	218.29 m ³ ·mol ⁻¹	Chemaxon
Polarizability	87.15 Å ³	Chemaxon
Number of Rings	7	Chemaxon
Bioavailability	0	Chemaxon
Rule of Five	No	Chemaxon
Ghose Filter	No	Chemaxon
Veber's Rule	No	Chemaxon
MDDR-like Rule	Yes	Chemaxon

Table 2.2 ADMET Properties of Paclitaxel

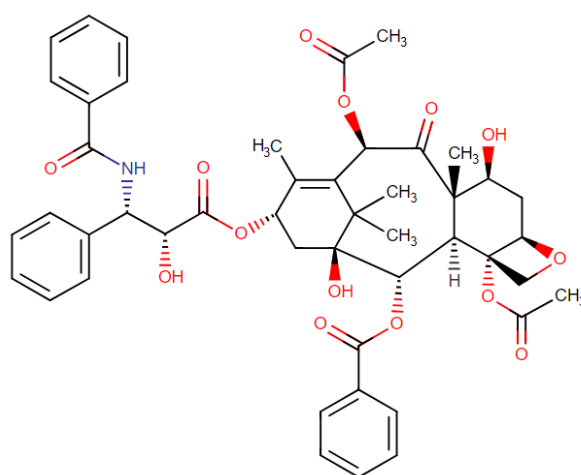


Figure 2.2 Structure of Paclitaxel

2.2.2 NAVITOCCLAX (DB12340):

Navitoclax has been used in trials studying the treatment and basic science of Solid Tumors, Non-Hodgkin's Lymphoma, EGFR Activating Mutation, Chronic Lymphoid Leukemia, and Hematological Malignancies, among others.

Navitoclax is an orally bioavailable small molecule inhibitor of Bcl-2 family proteins. It is a substance being studied in the treatment of lymphomas and other types of cancer. It blocks some of the enzymes that keep cancer cells from dying.

Navitoclax targets the Bcl-2 family of proteins, the major negative regulators of apoptosis. The Bcl-2 proteins, including Bcl-2, Bcl-xL, and Bcl-w, work by binding to two other groups of proteins-the executioners (Bax, Bak) that actually start the destruction pathway, and the sentinel proteins.

Cancer cells frequently overexpress the Bcl-2-like proteins, and thus, when they sustain DNA damage-from radiation, for example-they continue growing. Preventing the Bcl-2-like proteins from binding to the executioners might be able to trigger cell death in the tumor.

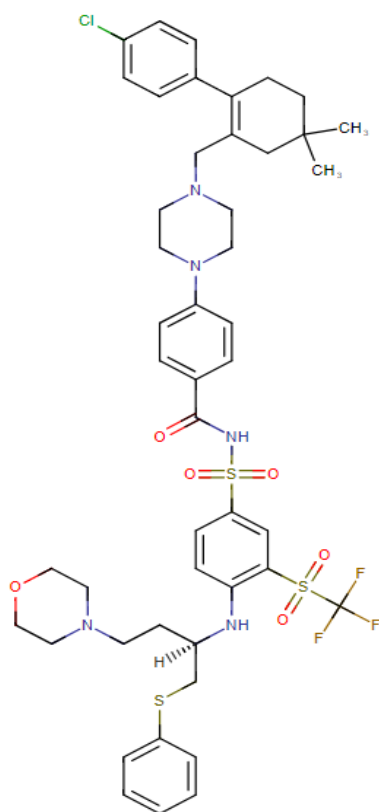
Chemical Formula of Navitoclax - $C_{47}H_{55}ClF_3N_5O_6S_3$ 

Figure 2.3 Structure of Navitoclax

PROPERTY	VALUE	SOURCE
Water Solubility	0.000212 mg/mL	<u>ALOGPS</u>
logP	7.77	<u>ALOGPS</u>
logP	7.93	<u>Chemaxon</u>
logS	-6.7	<u>ALOGPS</u>
pKa (Strongest Acidic)	4.26	<u>Chemaxon</u>
pKa (Strongest Basic)	8.3	<u>Chemaxon</u>
Physiological Charge	0	<u>Chemaxon</u>
Hydrogen Acceptor Count	10	<u>Chemaxon</u>
Hydrogen Donor Count	2	<u>Chemaxon</u>
Polar Surface Area	128.36 Å ²	<u>Chemaxon</u>
Rotatable Bond Count	16	<u>Chemaxon</u>
Refractivity	256.36 m ³ ·mol ⁻¹	<u>Chemaxon</u>
Polarizability	99.45 Å ³	<u>Chemaxon</u>
Number of Rings	7	<u>Chemaxon</u>
Bioavailability	0	<u>Chemaxon</u>
Rule of Five	No	<u>Chemaxon</u>
Ghose Filter	No	<u>Chemaxon</u>
Veber's Rule	No	<u>Chemaxon</u>
MDDR-like Rule	Yes	<u>Chemaxon</u>

Table 2.3 ADMET Properties of Navitoclax

2.2.3 TAK-901 (DB12756)

TAK-901 has been used in trials studying the treatment of Lymphoma, Myelofibrosis, Multiple Myeloma, Myeloid Metaplasia, and Advanced Solid Tumors, among others.

Chemical Formula of TAK 901 - $C_{28}H_{32}N_4O_3S$

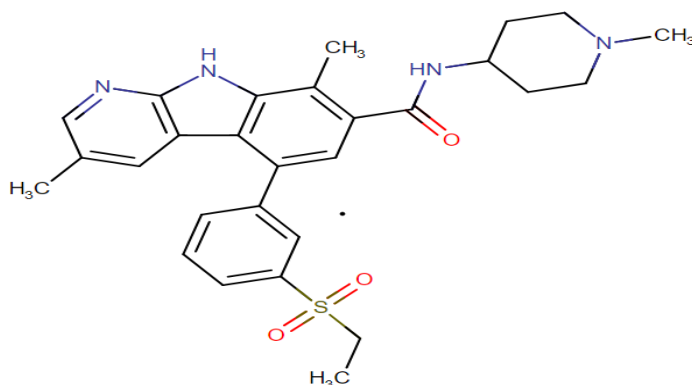


Figure 2.4 Structure of TAK-901

PROPERTY	VALUE	SOURCE
Water Solubility	0.00415 mg/mL	ALOGPS
logP	3.68	ALOGPS
logP	3.37	Chemaxon
logS	-5.1	ALOGPS
pKa (Strongest Acidic)	13.72	Chemaxon
pKa (Strongest Basic)	8.54	Chemaxon
Physiological Charge	1	Chemaxon
Hydrogen Acceptor Count	5	Chemaxon
Hydrogen Donor Count	2	Chemaxon
Polar Surface Area	95.16 Å ²	Chemaxon
Rotatable Bond Count	5	Chemaxon
Refractivity	144.58 m ³ ·mol ⁻¹	Chemaxon
Polarizability	56.38 Å ³	Chemaxon
Number of Rings	5	Chemaxon

Table 2.4 ADMET Properties of TAK-901

2.3 TARGET PROTEINS – ANTI CANCER ACTIVITY

🌟 P10415 · BCL2_HUMAN

Protein ⁱ	Apoptosis regulator Bcl-2	Amino acids	239 (go to sequence)
Gene ⁱ	BCL2	Protein existence ⁱ	Evidence at protein level
Status ⁱ	🌟 UniProtKB reviewed (Swiss-Prot)	Annotation score ⁱ	5/5
Organism ⁱ	Homo sapiens (Human)		

>sp|P10415|BCL2_HUMAN Apoptosis regulator Bcl-2 OS=Homo sapiens OX=9606 GN=BCL2 PE=1 SV=2

MAHAGRTGYDNREIVMKYIHYKLSQRGYEWDA GDVGAAPPGAAPAGIFSSQPGHTPHPA
ASRDPVARTSPLQTPAAPGAAAGPALSPVPPVVHLTLRQAGDDFSRRYRRDFAEMSSQLH
LTPFTARGRFATVVEELFRDGVNWGRIVAFFEFGGVMCVESVNREMSPLVDNIALWMTEY
LNRHLHTWIQDNGGWDAFVELYGPSMRPLFDFSWLSLKTLLSLALVGACITLGAYLGHK

🌟 Q07820 · MCL1_HUMAN

Protein ⁱ	Induced myeloid leukemia cell differentiation protein Mcl-1	Amino acids	350 (go to sequence)
Gene ⁱ	MCL1	Protein existence ⁱ	Evidence at protein level
Status ⁱ	🌟 UniProtKB reviewed (Swiss-Prot)	Annotation score ⁱ	5/5
Organism ⁱ	Homo sapiens (Human)		

>sp|Q07820|MCL1_HUMAN Induced myeloid leukemia cell differentiation protein Mcl-1 OS=Homo sapiens OX=9606 GN=MCL1 PE=1 SV=3

MFGLKRNAVIGLNL YCGGAGLGAGSGGATRPGGRLLATEKEASARREIGGGGEAGAVIGGS
AGASPPSTLTPDSRRVARPPPIGAEVPDVTATPARLLFFAPTRRAAPLEEMEAPAADAIM
SPEEELDGYEPEPLGKRPAVLPLELVGESGNNTSTDGSLPSTPPPAEEEEDELYRQSLE
IISRYLREQATGAKDTKPMGRSGATSRKALETLLRRVGDGVQRNHETAFQGMLRKLDIKNE
DDVKSLSRVMIHVFSDGVTNWGRIVTLISFGAFVAKHLKTINQESCIEPLAESITDVLVR
TKRDWLVKQRGWDGFEFFHVEDLEGGIRNVLLAFAGVAGVGAGLAYLIR

🌟 Q07817 · B2CL1_HUMAN

Protein ⁱ	Bcl-2-like protein 1	Amino acids	233 (go to sequence)
Gene ⁱ	BCL2L1	Protein existence ⁱ	Evidence at protein level
Status ⁱ	🌟 UniProtKB reviewed (Swiss-Prot)	Annotation score ⁱ	5/5
Organism ⁱ	Homo sapiens (Human)		

>sp|Q07817|B2CL1_HUMAN Bcl-2-like protein 1 OS=Homo sapiens OX=9606 GN=BCL2L1 PE=1 SV=1 MSQSNRELVDFLSYKLSQKGYWSQFSDVEENRTEAPEGTESEMETPSAINGNPSWHLA DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAFSDLTSQLHITPGTAY QSFEQVVNELFRDGVNWGRIVAFFSFGGALCVESVDKEMQVLVSRIAAMATYLNHLEP WIQENGWDTFVELYGNNAAESRKGQERFNRWFLTGMTVAGVVLLGSLFSRK

2.4 LIGAND SELECTION RULE

Ligand selection rules, often referred to in chemistry, particularly in the context of coordination complexes, dictate which ligands are likely to bind to a central metal ion based on certain criteria. These rules help predict the stability and geometry of coordination complexes.

1. Hard and Soft Acids and Bases (HSAB): This concept suggests that "hard" acids prefer to bind to "hard" bases, and "soft" acids prefer to bind to "soft" bases. Hard bases are typically small, highly charged ions (e.g., F^- , OH^-), while soft bases are larger, less electronegative atoms or molecules with lone pairs (e.g., I^- , CN^-). Hard acids are often found among the early transition metals and soft acids among the late transition metals.
2. Electronegativity: Ligands with higher electronegativity tend to be stronger donors and can form stronger bonds with the central metal ion.
3. Size and Steric Effects: Smaller ligands can form tighter bonds due to their ability to approach the metal ion more closely. Additionally, the steric hindrance caused by bulky ligands can influence the geometry of the resulting complex.
4. Charge: Ligands with a higher charge density (either positive or negative) tend to bind more strongly to the central metal ion.
5. Chelating Ligands: Ligands that can form multiple bonds (chelate) with the metal ion are often favored due to increased stability.
6. Field Theory: This theory, proposed by Alfred Werner, suggests that ligands can be classified as either weak-field or strong-field ligands based on their ability to cause splitting of d orbitals in the central metal ion. This leads to different electronic configurations and geometries in coordination complexes.

2.4.1 Lipinski's Rule of Five

Lipinski's Rule of Five, also known as the Pfizer's Rule of Five, is a set of guidelines in pharmacology used to evaluate the drug-likeness of chemical compounds, particularly those intended for oral administration. These rules were formulated by Christopher A. Lipinski in 1997 based on an analysis of compounds in the World Drug Index. The Rule of Five helps predict whether a chemical compound is likely to be orally bioavailable and pharmacologically active.

Molecular Weight (MW): The molecular weight of the compound should be less than 500 Daltons.

Lipophilicity (LogP): The octanol-water partition coefficient (LogP) should be less than 5. LogP is a measure of a compound's hydrophobicity or lipophilicity, indicating its ability to partition between octanol (lipophilic) and water (hydrophilic).

Hydrogen Bond Donors (HBD): The number of hydrogen bond donors should be less than 5. Hydrogen bond donors are atoms (usually nitrogen or oxygen) with hydrogen atoms attached that can potentially form hydrogen bonds.

Hydrogen Bond Acceptors (HBA): The number of hydrogen bond acceptors should be less than 10. Hydrogen bond acceptors are atoms (usually nitrogen or oxygen) capable of accepting hydrogen bonds due to lone pairs of electrons.

The rationale behind Lipinski's Rule of Five is that compounds that violate more than one of these rules are less likely to have favorable pharmacokinetic properties, such as absorption, distribution, metabolism, and excretion (ADME). However, it's essential to note that these rules are not absolute and are meant to serve as guidelines rather than strict thresholds. Many successful drugs exist outside of these parameters, and other factors, such as specific target binding and therapeutic index, also play crucial roles in drug development.

2.4.2 Veber's Rule

Veber's Rule, proposed by Chris Veber in 2002, is another set of guidelines used in drug design and optimization. It focuses on the molecular flexibility of compounds, particularly their rotatable bonds, and aims to predict the oral bioavailability of potential drug candidates. Veber's Rule suggests that compounds with a higher number of rotatable bonds may have

reduced oral bioavailability due to increased conformational flexibility, which can lead to poor absorption and metabolic instability.

Rotatable Bonds: The number of rotatable bonds in the molecule should be less than or equal to 10.

The rationale behind this rule is that compounds with a higher number of rotatable bonds are more flexible and can adopt a larger number of conformations, which may hinder their ability to bind to biological targets or may make them more susceptible to metabolic degradation.

Similar to Lipinski's Rule of Five, Veber's Rule serves as a guideline rather than a strict rule, and exceptions exist. However, it can be a useful tool in the early stages of drug discovery and optimization to prioritize compounds with a higher likelihood of oral bioavailability.

2.4.3 Ghose Filter

The Ghose Filter focuses on physicochemical properties of compounds and consists of the following criteria:

Molecular Weight (MW): The molecular weight of the compound should fall within the range of 160 to 480 Daltons.

Partition Coefficient (LogP): The octanol-water partition coefficient (LogP) should be between -0.4 and 5.6.

Hydrogen Bond Donors (HBD): The number of hydrogen bond donors should be between 0 and 5.

Hydrogen Bond Acceptors (HBA): The number of hydrogen bond acceptors should be between 0 and 10.

molar refractivity (MR): The molar refractivity should be between 40 and 130 cubic angstroms.

These criteria are based on statistical analyses of properties of known drugs. Compounds that fall within these ranges are considered to have a higher likelihood of being orally bioavailable and drug-like.

Similar to other rules and filters in drug design, the Ghose Filter is a guideline rather than a strict rule, and exceptions exist. However, it can be a useful tool in the early stages of drug discovery to prioritize compounds with favorable physicochemical properties.

2.4.4 MDDR-like Rule

The MDDR-like Rule, also known as the MDDR Rule or the MDL Drug Data Report Rule, is a set of guidelines used in drug design and optimization. It was developed based on an analysis of compounds in the MDL Drug Data Report (MDDR) database, which contains information on small molecule drugs.

The MDDR-like Rule focuses on several physicochemical properties of compounds and aims to identify molecules that are similar to known drugs in terms of their chemical characteristics. This similarity suggests that these compounds may have a higher likelihood of success in drug development. While the specific criteria for the MDDR-like Rule may vary slightly depending on the source, they typically include the following:

Molecular Weight (MW): The molecular weight of the compound should fall within a certain range, typically between 150 and 500 Daltons.

LogP (Partition Coefficient): The octanol-water partition coefficient (LogP) should be within a specific range, indicating moderate lipophilicity.

Hydrogen Bond Donors (HBD): The number of hydrogen bond donors should be limited, typically fewer than 5.

Hydrogen Bond Acceptors (HBA): The number of hydrogen bond acceptors should also be limited, typically fewer than 10.

Aromatic Rings: The number of aromatic rings in the molecule should be within a certain range, often less than 6.

Polar Surface Area (PSA): The polar surface area of the compound should be moderate, typically less than 140 square angstroms.

The rationale behind the MDDR-like Rule is that compounds with properties similar to those of known drugs are more likely to exhibit favorable pharmacokinetic and pharmacodynamic properties, making them promising candidates for drug development.

As with other drug-likeness rules and filters, the MDDR-like Rule is not absolute, and exceptions exist. However, it can be a useful tool in drug discovery to prioritize compounds with characteristics associated with successful drugs.

2.5 PROTEIN -LIGAND DOCKING

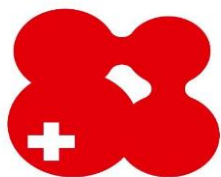
- Protein–ligand docking aims to predict the position and orientation of a ligand when bound to a protein receptor or enzyme. They are used for virtual screening of large databases of chemicals to select likely drug candidates.

- Significance of Computational Ability
- Advancements in computational ability, such as with AlphaFold, are enhancing the accuracy of protein structure predictions and protein-ligand docking.
- Protein-ligand docking is a computational technique used to predict the binding of a small molecule (ligand) to a target protein.
- Identifying the binding site on the target protein. This can be done either by using known information about the binding site, or by using computational methods to predict potential binding pockets on the protein surface.
- Generating possible conformations (poses) of the ligand within the binding site. This involves sampling the six degrees of freedom (three translational and three rotational) of the ligand as well as any conformational flexibility of the ligand itself.
- Evaluating and ranking the generated poses using scoring functions that estimate the binding affinity between the ligand and protein. The pose with the best (lowest) score is considered the predicted binding mode.
- Protein-ligand docking is widely used in structure-based drug discovery to identify potential drug candidates that can bind to and modulate the activity of target proteins associated with diseases. It allows rapid virtual screening of large chemical libraries to find hits that can then be further optimized through medicinal chemistry. The key advantages of protein-ligand docking are its computational efficiency and ability to provide atomic-level details of the binding interaction. However, limitations include the need to treat the protein as rigid in most cases, and the imperfect accuracy of current scoring functions. Ongoing research aims to improve docking methods to better account for protein and ligand flexibility as well as develop more accurate scoring functions.

CHAPTER 3

MATERIALS AND METHODS

3.1 SWISS MODEL



SWISS-MODEL is a web-based integrated service dedicated to protein structure homology modelling. It guides the user in building protein homology models at different levels of complexity.

Building a homology model comprises four main steps: (i) identification of structural template(s), (ii) alignment of target sequence and template structure(s), (iii) model-building, and (iv) model quality evaluation. These steps require specialised software and integrate up-to-date protein sequence and structure databases. Each of the above steps can be repeated interactively until a satisfying modelling result is achieved.

3.1.1 Workplace

The SWISS-MODEL Workplace integrates programs and databases required for protein structure modelling in a web-based workspace. Depending on the complexity of the modelling task, different modes of usage can be applied, in which the user has different levels of control over individual modelling steps: automated mode, alignment mode, and project mode. A fully automated mode is used when a sufficiently high sequence identity between target and template (>50%) allows for no human intervention at all.

In this case only the sequence or UniProt accession code of the protein is required as input. The alignment mode enables the user to input their own target-template alignments from which the modelling procedure starts (i.e. search for templates step is skipped and rarely only minor changes in the provided alignment are made). The project mode is used in more difficult cases, when manual corrections of target-template alignments are needed to improve the quality of the resulting model. In this mode the input is a project file that can be generated by the DeepView (Swiss Pdb Viewer) visualization and structural analysis tool, allow the user to examine and manipulate the target-template alignment in its structural context.

In all three cases the output is a pdb file with atom coordinates of the model or a DeepView project file. The four main steps of homology modelling may be repeated iteratively until a satisfactory model is achieved.

The SWISS-MODEL Workspace is accessible via the Expasy web server, or it can be used as part of the program DeepView (Swiss Pdb-Viewer). As of September 2015, has been cited 20000 times in scientific literature, making it one of the most widely used tools for protein structure modelling. The tool is free for academic use.

3.1.2 Model Building

Models are computed by the SWISS-MODEL server homology modelling pipeline (Waterhouse et.al) which relies on ProMod3, an in-house comparative modelling engine based on OpenStructure.

ProMod3 extracts initial structural information from the template structure. Insertions and deletions, as defined by the sequence alignment, are resolved by first searching for viable candidates in a structural database. Final candidates are then selected using statistical potentials of mean force scoring methods. If no candidates can be found, a conformational space search is performed using Monte Carlo techniques. Non-conserved side chains are modelled using an in-house backbone-dependent rotamer library. The optimal configuration of rotamers is estimated using the graph-based TreePack algorithm (*Xu et al.*) by minimising the SCWRL4 energy function (*Krivov et al.*). As a final step, small structural distortions, unfavourable interactions or clashes introduced during the modelling process are resolved by energy minimisation. ProMod3 uses the OpenMM library (Eastman et al) to perform the computations and the CHARMM22/CMAP force field for parameterisation.

3.1.3 Model Evaluation

GMQE and QMEANDisCo global give an overall model quality measurement between 0 and 1, with higher numbers indicating higher expected quality. GMQE is coverage dependent, i.e. a model covering only half of the target sequence is unlikely to get a score above 0.5. QMEANDisCo on the other hand evaluates the model 'as is' without explicit coverage dependency.

GMQE (Global Model Quality Estimate) is a quality estimate which combines properties from the target-template alignment and the template structure. They are combined using a

multilayer perceptron trained to predict the IDDT score of the resulting model. The GMQE is available before building an actual model and thus helpful in selecting optimal templates for the modelling problem at hand. Once a model is built, the GMQE gets updated for this specific case by also taking into account the QMEANDisCo global score of the obtained model in order to increase reliability of the quality estimation. If a template structure originates from AlphaFold DB, GMQE is a heuristic that sums per-residue PLDDT values of aligned template residues and normalizes by target sequence length. Once a model is built, the GMQE is updated and represents the summed per-residue quality estimates normalized by target sequence length. Again, per-residue quality estimates are estimated with an AlphaFold DB specific heuristic.

QMEANDisCo global score is the average per-residue QMEANDisCo score (see below) which has been found to correlate well with the IDDT score. The provided error estimate is based on QMEANDisCo global scores estimated for a large set of models and represents the root mean squared difference between QMEANDisCo global score and IDDT (the ground truth). As the reliability of the prediction depends on model size, the provided error estimate is calculated based on models of similar size to the input. The QMEANDisCo global score is not computed for models that use AlphaFold DB templates.

QMEAN Z-score analysis is deprecated and the GMQE and QMEANDisCo global scores should be consulted for global model quality estimates instead. It is based on 4 statistical potentials of mean force and their linear combination: the "QMEAN" score. All scores, 5 in total, are compared with what one would expect from experimentally determined structures of similar size using Z-scores. In other words: How many standard deviations from the mean is my model score given a score distribution from a large set of experimentally determined structures. Z-scores around 0.0 therefore reflect a "native-like" structure and, as a rule of thumb, a "QMEAN" Z-score below -4.0 indicates a model with low quality. This is illustrated by the "Comparison" plot the x-axis shows protein length (number of residues). The y-axis is the "QMEAN" score. Every dot represents one experimental protein structure. Black dots are experimental structures with a "QMEAN" score within 1 standard deviation of the mean ($|Z\text{-score}|$ between 0 and 1), experimental structures with a $|Z\text{-score}|$ between 1 and 2 are grey. Experimental structure that are even further from the mean are light grey.

The actual model is represented as a red star. The QMEAN Z-score analysis is not computed for models that use AlphaFold DB templates.

3.1.4 Model Report

The SWISS-MODEL Homology Modelling Report offers a summary of all Models built in the project.

Note: The report is accessible (i) per model via a drop-down menu, next to the model in the Models view or (ii) for all models in report.html in the downloaded file when choosing to download the project by pressing the download button below the project title.

It is structured in the following sections:

- **Model building Report:** Contains project name, project date and references. The target sequence is in Table T1 of the Report.
- **Results:** Version of the SWISS-MODEL template library and PDB release. All identified templates are listed in Table T2.
- **Models:** Models are listed sequentially with each entry showing a picture of the model, a link to the PDB file, the version of the modelling engine, the oligomeric state, the ligands (if any), the global model quality estimate, and the QMEAN score. A graphical representation of the QMEAN score and its four terms separately, the local quality estimate plot, and the comparison with non-redundant set of PDB structures are also provided. For the template, a link to the template itself is provided together with the following information: the title of the structure, the target sequence coverage, the sequence identity to the target, the experimental method used to obtain the structure (and the resolution, if applicable), the oligomeric state, the ligands (if any), the sequence similarity to the target, the template search method used.
- **Save Project Locally:** Allows to download the project as a zip file. The main folder contains the Model report images folder (banner for the Report), and the model folder. Each model has its own subfolder.

3.2 DRUGBANK

DrugBank Online is a comprehensive, free-to-access, online database containing information on drugs and drug targets. As both a bioinformatics and a cheminformatics resource, we combine detailed drug (i.e. chemical, pharmacological and pharmaceutical) data with comprehensive drug target (i.e. sequence, structure, and pathway) information. DrugBank Online is widely used by the drug industry, medicinal chemists, pharmacists,

physicians, students and the general public. Because of its broad scope, comprehensive referencing, and detailed data descriptions, DrugBank is enabling major advancements across the data-driven medicine industry.

DrugBank Online is offered to the public as a free-to-access resource. Use and re-distribution of the content of DrugBank Online or the DrugBank Data, in whole or in part, for any purpose requires a license. Academic users may apply for a free license for certain use cases and all other users require a paid license.



3.3 AUTODOCK VINA

AutoDock Vina is an open-source program for doing molecular docking. It was originally designed and implemented by Dr. Oleg Trott in the Molecular Graphics Lab (now CCSC) at The Scripps Research Institute.

3.3.1 Features

AutoDock Vina significantly improves the average accuracy of the binding mode predictions compared to AutoDock 4, judging by our tests on the training set used in AutoDock 4 development. Additionally, and independently, AutoDock Vina has been tested against a virtual screening benchmark called the director useful of decoys by the Watowich groups and was found to be “a strong competitor against the other programs, and at the top of the pack in many cases”. It should be noted that all six of the other docking programs, to which it was compared, are distributed commercially.

A) AutoDock Tools Compatibility

For its input and output, Vina uses the same PDBQT molecular structure file format used by Autodock. PDBQT files can be generated (interactively or in batch mode) and viewed using MGL Tools. Other files, such as the AutoDock and AutoGrid parameter files (GPF, DPF) and grid map files are not needed.

Binding mode prediction accuracy on the test set. “AutoDock” refers to AutoDock 4, and “Vina” to AutoDock Vina 1. Vina’s design philosophy is not to require the user to understand its implementation details, tweak obscure search parameters, cluster results or know

advanced algebra (quaternions). All that is required is the structures of the molecules being docked and the specification of the search space including the binding site.

Calculating grid maps and assigning atom charges is not needed. The usage summary can be printed with “vina --help “. The summary automatically remains in sync with the possible usage scenarios.

B) Implementation Quality

- By design, the results should not have a statistical bias related to the conformation of the input structure.
- Attention is paid to checking the syntactic correctness of the input and reporting errors to the user in a lucid manner.
- The invariance of the covalent bond lengths is automatically verified in the output structures.
- Vina avoids imposing artificial restrictions, such as the number of atoms in the input, the number of torsions, the size of the search space, the exhaustiveness of the search, etc.

C) Flexible Side Chains: Like in AutoDock 4, some receptor side chains can be chosen to be treated as flexible during docking.

3.4 METHODOLOGY - AutoDock Vina

3.4.1 Coordinate Preparation with AutoDock Tools

STEP 1

Generate the ligand coordinate file. A coordinate set that includes hydrogen atoms is required. This may be obtained in a variety of ways, including experimental coordinates from the Protein Data Bank (www.pdb.org) or Cambridge Crystallographic Database (ccdc.cam.ac.uk), or structure generation methods such as the CACTUS server (cactus.nci.nih.gov/translate). The file 1iep_ligandH.pdb is provided for use as a tutorial for this protocol (all example files are supplied in [Supplementary Data](#)) it includes ligand coordinates taken from PDB entry 1iep, to which all hydrogen atoms have been added and manually adjusted to the known protonation state. Start AutoDockTools (ADT) ([Figure 1](#)) and set the working directory by clicking “File->Preferences->Set”.

Type your working directory path name into the “Startup Directory” box and click “Set”. Click “Dismiss” at the bottom of the window.

STEP 2

Read the atomic coordinates. To do this, first select “Ligand->Input->Open” and use the “Files of type” menu to choose “PDB files”. Click on your coordinate file, in this case, 1iep_ligandH.pdb, and click “Open”. ADT will read the coordinates, add charges if necessary, merge non-polar hydrogens, and assign appropriate atom types. At this point, the ligand will be displayed in the viewer window, with aromatic carbons in green. Click “OK” on the popup to continue.

CRITICAL STEP *If your coordinate set does not include hydrogen positions, click “File->Read Molecule” and choose your coordinate file. Then, by clicking “Edit->Hydrogens->Add”, you will add all hydrogens by default. Select “Ligand->Input->Choose” to choose the ligand molecule.*

STEP 3

Prepare a PDBQT file by selecting “Ligand->TorsionTree->DetectRoot”, this will define the center of the torsion tree. Selecting “Ligand->TorsionTree->ChooseTorsions” will launch a window that allows choice of torsional degrees of freedom. Rotatable bonds are in green, rigid portions are in red, and potentially rotatable bonds that are currently set as “not rotatable” (such as the peptide bond at the center of imatinib) are in magenta. Clicking on bonds will switch the rotation flexibility on and off. When finished, click “Done”. Click “Ligand->Output->SaveAsPDBQT”, then select “Save” to write the file 1iep_ligandH.pdbqt.

STEP 4

Generate the receptor coordinate file. A receptor coordinate file with all hydrogen atoms is required. If you are using experimental structures (for instance, from the PDB), use a text editor to remove water, ligands, cofactors, ions, etc. that should not be included in the receptor. The file 1iep_receptorH.pdb is provided for use as a tutorial for this protocol and it includes receptor coordinates taken from PDB entry 1iep. Open the file by selecting “Grid->Macromolecule->Open”, use the “Files of type” menu to choose “all files”. Click on your coordinate file, in this case, 1iep_receptorH.pdb, and click “Open”. ADT will read coordinates, add charges, merge non-polar hydrogens, and assign appropriate atom types. Click “OK” to accept the changes. A window will pop up to write the PDBQT file. Click “Save” to write a file 1iep_receptorH.pdbqt

CRITICAL STEP If your coordinate set does not include hydrogen positions, Click “File->Read Molecule” and choose your coordinate file. Then, selecting “Edit->Hydrogens->Add” will add all hydrogens by default. Click “Grid->Macromolecule->Choose” to choose the receptor molecule.

3.4.2 Virtual Screening with Raccoon2 and AutoDock Vina

STEP 1

Start Raccoon2 and configure the server. Raccoon is designed to run virtual screening on a large computational resource, such as a Linux cluster. When you add a new server, you must configure the connection and install one or more docking services. Launch Raccoon2 and create a new server connection by clicking the three-gear icon to open the Connection Manager. Several options need to be set: server name (a name to identify the resource); address (the host name or IP address of the server); and a username and password. Click “Save” and close the Connection Manager.

Install a docking service by selecting the server in the Connection menu. Click on the raccoon-paw icon to prepare (“raccoonize”) the server. Click on the “Service Manager” icon (a wand) and add a new service by clicking on the plus icon, and give it a descriptive name of your choice (“Vina docking service”). Click on “Install AutoDock Vina on the server”, and accept several default selections. Click Save, and close the Service Manager. Finally, from the “Available Services” panel, select the “Vina docking service”

STEP 2

Set up the Ligand Library. This protocol describes virtual screening of C-Abl with a small library of 500 compounds that includes the known inhibitor imatinib. Tools for generating and processing ligand libraries are available at <http://autodock.scripps.edu/resources/databases>. Ligand libraries are stored on the server and made available for docking jobs, to reduce redundancy and allow tracing and reproducing experiments.

In the LIGANDS tab, select “AddLigand->ScanDirectory” and browse to the directory containing the ligand library. Close the report window and select “Upload”, type a descriptive library name of your choice and click “Start.” Close after upload is complete, close the Library Manager. Right click to select the desired ligand library.

STEP 3

Set up the receptor coordinates. In the RECEPTORS tab, select “Add->ScanDirectory” browse to the receptor PDBQT file (created in Step 4), and click “OK”. Close the report window.

STEP 4

Configure AutoDock Vina docking parameters. In the CONFIG tab: load a receptor structure in the 3D viewer by selecting files in the Receptor list. Use “Center” and “Size” thumbwheels to define the search space, or place the mouse cursor over boxes and type numeric values. Modify docking parameters if desired or load the config.txt file created.

STEP 5

Perform the virtual screening calculation. In the JOB MANAGER tab, ensure that all “Cluster submission requirements” are green, then click “Submit.” Several options will need to be specified: set “Project” to “new” and type a descriptive name of your choice; set “Experiment” to “new” and type a descriptive name of your choice; and click “OK” to start the calculation. To follow the status of the docking, right click on the experiment name and select “update status.” When one or more jobs in the experiment are finished, download them individually by right-clicking on the job name, or together by right-clicking on the experiment and select “Download results.”

STEP 6

Filter and analyze the results. To assist with filtering and selection, Raccoon calculates docking pose properties such as interaction, score, and ligand efficiency. In the ANALYSIS->DATA SOURCE tab, click on “Add results->Select downloaded results”, and browse to the results directory. Choose the receptor file to be used for processing results and type a descriptive log file name or select the config.txt file generated.

The filter panel allows a variety of filters to be applied to the results (“Energy” will show poses within an energy range; “Ligand Efficiency” will select poses within a range of ligand efficiency; and with “Target Interactions,” use “+” to filter by selected interactions with receptor atoms.

CHAPTER 4

RESULTS AND DISCUSSION

P10415 · BCL2_HUMAN

Protein ⁱ	Apoptosis regulator Bcl-2	Amino acids	239 (go to sequence)
Gene ⁱ	BCL2	Protein existence ⁱ	Evidence at protein level
Status ⁱ	 UniProtKB reviewed (Swiss-Prot)	Annotation score ⁱ	
Organism ⁱ	Homo sapiens (Human)		

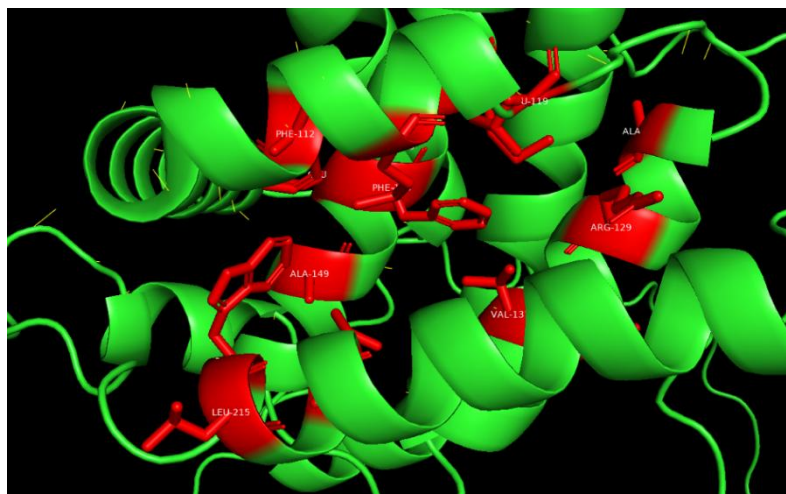


Figure 4.1 Target Protein BCL-2 with its Active Sites (Red) (Petros *et al.*, 2001)

4.1 Active site amino acids

ALA126_ALA149_ARG129_GLU152_LEU119_LEU215_LEU217_MET115_MET157_PHE
112_PHE153_THR132_TRP214_VAL133

4.2 Docking Score

DrugBank ID	Drug Name	Affinity (kcal/mol)
DB06077	Lumateperone	-8
DB12978	Pexidartinib	-8
DB08896	Regorafenib	-7.9
DB14075	Imidurea	-7.9
DB00398	Sorafenib	-7.8
DB00734	Risperidone	-7.8
DB06684	Vilazodone	-7.8
DB09280	Lumacaftor	-7.8
DB01184	Domperidone	-7.7
DB06210	Eltrombopag	-7.7

4.3 DISCUSSION

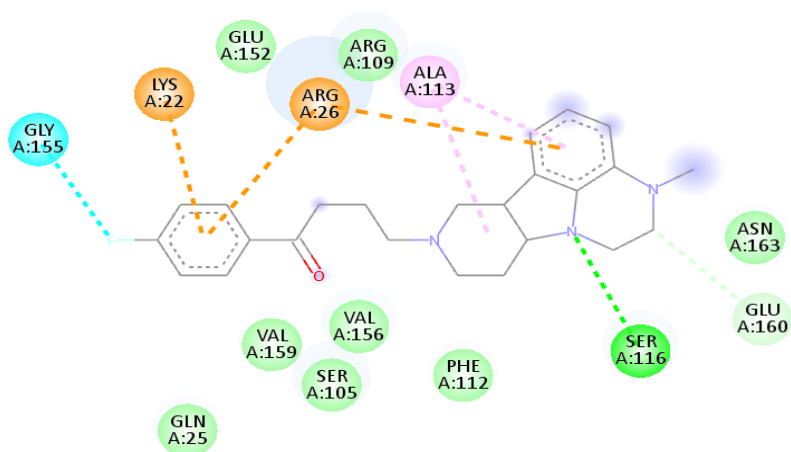


Figure 4.2 Docking Visualisation of Lumateperone to BCL2

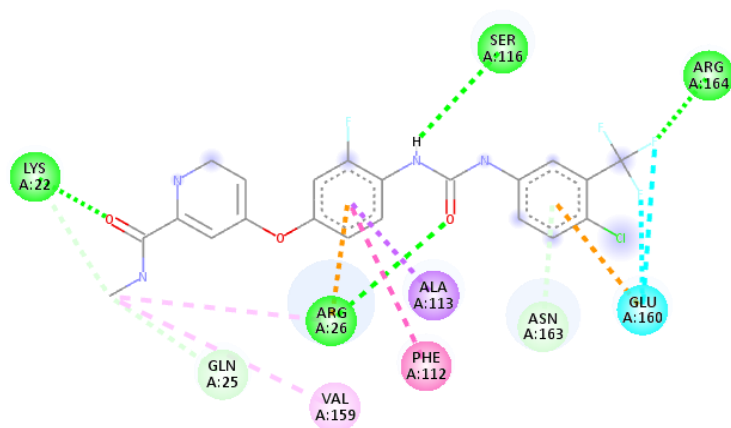


Figure 4.3 Docking Visualisation of Pexidartinib to BCL2

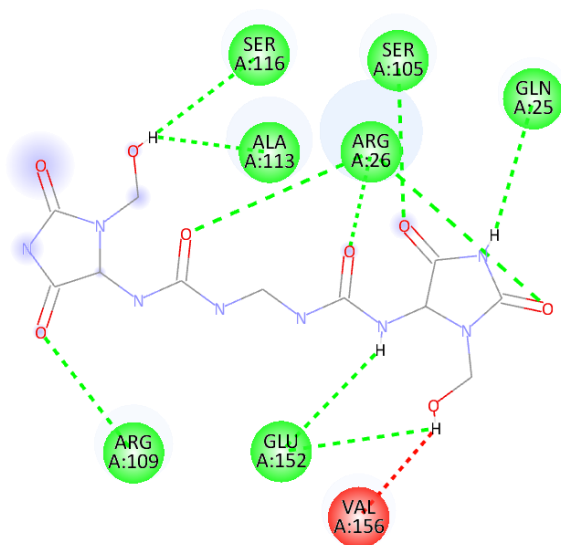


Figure 4.4 Docking Visualisation of Regorafenib to BCL2

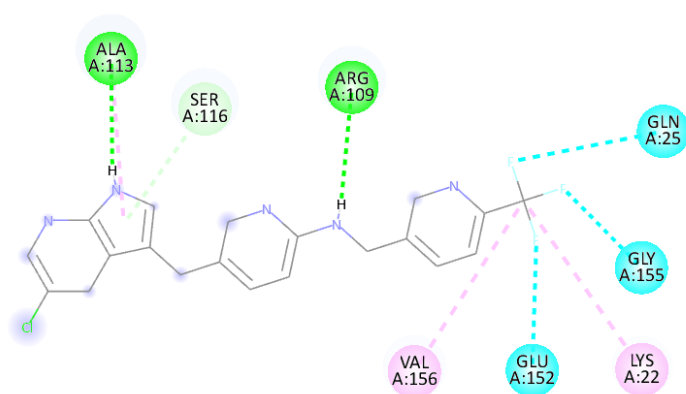


Figure 4.5 Docking Visualisation of Imidurea to BCL2



Figure 4.6 Inference Pictograph

The results of our in-silico analyses reveal promising interactions between repurposed drugs and various members of the BCL protein family. Notably, certain drugs like **lumateperone** and **pexidartinib** exhibit high binding affinities and favorable pharmacokinetic profiles, suggesting their potential as effective modulators of BCL proteins in cancer and neurodegenerative disorders. Molecular dynamics simulations elucidate the stability of drug-protein complexes and highlight key molecular interactions governing their binding. Virtual screening identifies additional repurposed drugs with favorable drug-like properties, expanding the pool of potential candidates for further investigation. **Pexidartinib** is an antitumor agent that is used for the treatment of rare disease tenosynovial giant cell tumors (TGCT) by inhibiting colony-stimulating factor 1 and its receptor. TGCT is a rare, non-malignant neoplasm that causes abnormal growth and damage to the synovium, bursae, or tendon sheaths. Recruitment of immune cells, specifically macrophages, is closely associated with the tumor mass formation in TGCTs. The recruitment, proliferation, and irreversible differentiation of macrophages are regulated by colony-stimulating factor-1 (CSF-1), which is a cytokine that is often translocated and highly expressed in TGCTs. Pexidartinib targets the CSF1/CSF1R pathway as a selective CSF1R inhibitor.

Lumateperone is a novel 2nd generation antipsychotic used to manage both positive and negative symptoms in patients with schizophrenia. Dopamine abnormalities, specifically in the prefrontal and mesolimbic brain regions, are consistent in people with schizophrenia. In addition to dopamine, other neurotransmitters such as serotonin, glutamate, GABA and acetylcholine are thought to play a role. Lumateperone is unique among second generation antipsychotics based on its target profile and dopamine D2 receptor occupancy. Unlike other antipsychotics, lumateperone has partial agonist activity at presynaptic dopamine (D2) receptors, resulting in reduced presynaptic release of dopamine, and antagonistic activity at postsynaptic dopamine (D2) receptors. These characteristics allow lumateperone to efficiently reduce dopamine signaling. These 2 drugs Pexidartinib and Lumateperone have exceptionally high binding affinity to our BCL2 protein target. They can potentially be used as a repurposed drug for cancer.

CHAPTER 5

SUMMARY AND CONCLUSION

Our in-silico investigation identified several existing drugs with high binding affinities to specific members of the BCL protein family implicated in cancer based on Lipinski's rule of Five, i.e., The molecular weight of the compound should be less than 500 Daltons. The octanol-water partition coefficient (LogP) should be less than 5. LogP is a measure of a compound's hydrophobicity or lipophilicity, indicating its ability to partition between octanol (lipophilic) and water (hydrophilic). The number of hydrogen bond donors should be less than 5. Hydrogen bond donors are atoms (usually nitrogen or oxygen) with hydrogen atoms attached that can potentially form hydrogen bonds. The number of hydrogen bond acceptors should be less than 10. Hydrogen bond acceptors are atoms (usually nitrogen or oxygen) capable of accepting hydrogen bonds due to lone pairs of electrons from 2619 approved drugs from DrugBank. Among these, compounds such as Lumateperone, Pexidartinib, Regorafenib, Imidurea, Sorafenib, Risperidone, Vilazodone, Lumacaftor, Domperidone and Eltrombopag exhibited strong interactions with key BCL proteins, suggesting their potential for repurposing as cancer therapeutics. Of them, Lumateperone and Pexidartinib has the highest Binding Affinity to our target BCL2 protein. **Lumateperone** is a novel 2nd generation antipsychotic used to manage both positive and negative symptoms in patients with schizophrenia. **Pexidartinib** is an antitumor agent that is used for the treatment of rare disease tenosynovial giant cell tumors (TGCT) by inhibiting colony-stimulating factor 1 and its receptor. Molecular dynamics simulations provided insights into the stability and dynamics of drug-protein complexes, elucidating the structural basis of their binding. These findings highlight the feasibility of repurposing existing drugs to target the BCL protein family for cancer treatment, paving the way for further experimental validation and clinical translation.

CHAPTER 6

REFERENCES:

1. Al-Zabeeby, A., Vogler, M., Milani, M., Richards, C., Alotibi, A., Greaves, G., ... & Varadarajan, S. (2018). Targeting intermediary metabolism enhances the efficacy of bh3 mimetic therapy in hematologic malignancies. *Haematologica*, 104(5), 1016-1025. <https://doi.org/10.3324/haematol.2018.204701>
2. Apte, S. S., Mattei, M.-G., & Olsen, B. R. (1995). Mapping of the human BAX gene to chromosome 19q13.3–q13.4 and isolation of a novel alternatively spliced transcript, BAX δ . *Genomics*, 26(3), 592–594. [https://doi.org/10.1016/0888-7543\(95\)80180-t](https://doi.org/10.1016/0888-7543(95)80180-t)
3. Biasini, M., Schmidt, T., Bienert, S., Mariani, V., Studer, G., Haas, J., Johner, N., Schenk, A. D., Philippsen, A., & Schwede, T. (2013). OpenStructure: An integrated software framework for computational structural biology. *Acta Crystallographica Section D Biological Crystallography*, 69(5), 701–709. <https://doi.org/10.1107/s0907444913007051>
4. Cava, C., Bertoli, G., & Castiglioni, I. (2018). In silico identification of drug target pathways in breast cancer subtypes using pathway cross-talk inhibition. *Journal of Translational Medicine*, 16(1). <https://doi.org/10.1186/s12967-018-1535-2>
5. Chen, T., Shu, X., Zhou, H., Beckford, F. A., & Misir, M. (2023). Algorithm selection for protein–ligand docking: Strategies and analysis on ACE. *Scientific Reports*, 13(1). <https://doi.org/10.1038/s41598-023-35132-5>
6. Cova, T., Bento, D., & Nunes, S. (2019). Computational approaches in theranostics: mining and predicting cancer data. *Pharmaceutics*, 11(3), 119. <https://doi.org/10.3390/pharmaceutics11030119>
7. Czabotar, P., Lessène, G., Strasser, A., & Adams, J. (2013). Control of apoptosis by the bcl-2 protein family: implications for physiology and therapy. *Nature Reviews Molecular Cell Biology*, 15(1), 49-63. <https://doi.org/10.1038/nrm3722>
8. Delft, M., Wei, A., Mason, K., Vandenberg, C., Chen, L., Czabotar, P., ... & Huang, D. (2006). The bh3 mimetic abt-737 targets selective bcl-2 proteins and efficiently induces apoptosis via bak/bax if mcl-1 is neutralized. *Cancer Cell*, 10(5), 389-399. <https://doi.org/10.1016/j.ccr.2006.08.027>
9. Eastman, P., Swails, J., Chodera, J. D., McGibbon, R. T., Zhao, Y., Beauchamp, K. A., Wang, L.-P., Simmonett, A. C., Harrigan, M. P., Stern, C. D., Wiewiora, R. P., Brooks, B. R., & Pande, V. S. (2017). OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLOS Computational Biology*, 13(7), e1005659. <https://doi.org/10.1371/journal.pcbi.1005659>
10. Eberhardt, J., Santos-Martins, D., Tillack, A., & Forli, S. (2021). *AutoDock Vina 1.2.0: New*

- docking methods, expanded force field, and python bindings. American Chemical Society (ACS). <http://dx.doi.org/10.26434/chemrxiv.14774223>
11. Eliseev, R. A., Malecki, J., Lester, T., Zhang, Y., Humphrey, J., & Gunter, T. E. (2009). Cyclophilin D interacts with bcl2 and exerts an anti-apoptotic effect. *Journal of Biological Chemistry*, 284(15), 9692–9699. <https://doi.org/10.1074/jbc.m808750200>
 12. Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., & Olson, A. J. (2016). Computational protein–ligand docking and virtual drug screening with the AutoDock suite. *Nature Protocols*, 11(5), 905–919. <https://doi.org/10.1038/nprot.2016.051>
 13. Gaitanis, A., & Staal, S. (2010). Liposomal Doxorubicin and nab-Paclitaxel: Nanoparticle Cancer Chemotherapy in Current Clinical Use. In *Methods in Molecular Biology* (pp. 385–392). Humana Press. http://dx.doi.org/10.1007/978-1-60761-609-2_26
 14. Gong, Q. (2023). B-cell lymphoma-2 family proteins in the crosshairs: small molecule inhibitors and activators for cancer therapy. *Medicinal Research Reviews*, 44(2), 707-737. <https://doi.org/10.1002/med.21999>
 15. Gustafsson, Å. B., Tsai, J. G., Logue, S. E., Crow, M. T., & Gottlieb, R. A. (2004). Apoptosis repressor with caspase recruitment domain protects against cell death by interfering with bax activation. *Journal of Biological Chemistry*, 279(20), 21233–21238. <https://doi.org/10.1074/jbc.m400695200>
 16. Hamis, S., Powathil, G., & Chaplain, M. (2019). Blackboard to bedside: a mathematical modeling bottom-up approach toward personalized cancer treatments. *Jco Clinical Cancer Informatics*, (3), 1-11. <https://doi.org/10.1200/cci.18.00068>
 17. Hashemi, Z., Zarei, M., Fath, M., Ganji, M., Farahani, M., Afsharnouri, F., ... & Khalili, S. (2021). In silico approaches for the design and optimization of interfering peptides against protein–protein interactions. *Frontiers in Molecular Biosciences*, 8. <https://doi.org/10.3389/fmolb.2021.669431>
 18. Hoffman-Luca, C. G., Ziazadeh, D., McEachern, D., Zhao, Y., Sun, W., Debussche, L., & Wang, S. (2015). Elucidation of acquired resistance to bcl-2 and MDM2 inhibitors in acute leukemia in vitro and in vivo. *Clinical Cancer Research*, 21(11), 2558–2568. <https://doi.org/10.1158/1078-0432.ccr-14-2506>
 19. Kang, M. and Reynolds, C. (2009). Bcl-2 inhibitors: targeting mitochondrial apoptotic pathways in cancer therapy. *Clinical Cancer Research*, 15(4), 1126-1132. <https://doi.org/10.1158/1078-0432.ccr-08-0144>
 20. Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., Neveu, V., Djoumbou, Y., Eisner, R., Guo, A. C., & Wishart, D. S. (2010). DrugBank 3.0: A comprehensive resource for “Omics” research on drugs. *Nucleic Acids Research*, 39(Database), D1035–D1041. <https://doi.org/10.1093/nar/gkq1126>
 21. Krivov, G. G., Shapovalov, M. V., & Dunbrack, R. L., Jr. (2009). Improved prediction of protein

- side-chain conformations with SCWRL4. *Proteins: Structure, Function, and Bioinformatics*, 77(4), 778–795. <https://doi.org/10.1002/prot.22488>
22. Law, V., Knox, C., Djoumbou, Y., Jewison, T., Guo, A. C., Liu, Y., Maciejewski, A., Arndt, D., Wilson, M., Neveu, V., Tang, A., Gabriel, G., Ly, C., Adamjee, S., Dame, Z. T., Han, B., Zhou, Y., & Wishart, D. S. (2013). DrugBank 4.0: Shedding new light on drug metabolism. *Nucleic Acids Research*, 42(D1), D1091–D1097. <https://doi.org/10.1093/nar/gkt1068>
 23. Lehmann, T., Miskiewicz, J., Szostak, N., Szachniuk, M., Grodecka-Gazdecka, S., & Jagodziński, P. (2020). In vitro and in silico analysis of mir-125a with rs12976445 polymorphism in breast cancer patients. *Applied Sciences*, 10(20), 7275. <https://doi.org/10.3390/app10207275>
 24. Liu, Q., Moldoveanu, T., Sprules, T., Matta-Camacho, E., Mansur-Azzam, N., & Gehring, K. (2009). *Solution Structure of human Mcl-1 complexed with human Bid_BH3 peptide*. <https://doi.org/10.2210/pdb2kbw/pdb>
 25. Lock, R., Carol, H., Houghton, P. J., Morton, C. L., Kolb, E. A., Gorlick, R., Reynolds, C. P., Maris, J. M., Keir, S. T., Wu, J., & Smith, M. A. (2007). Initial testing (stage 1) of the BH3 mimetic ABT-263 by the pediatric preclinical testing program. *Pediatric Blood & Cancer*, 50(6), 1181–1189. <https://doi.org/10.1002/pbc.21433>
 26. Masood, A., Azmi, A., & Mohammad, R. (2011). Small molecule inhibitors of bcl-2 family proteins for pancreatic cancer therapy. *Cancers*, 3(2), 1527-1549. <https://doi.org/10.3390/cancers3021527>
 27. Meng, X.-Y., Zhang, H.-X., Mezei, M., & Cui, M. (2011a). Molecular docking: A powerful approach for structure-based drug discovery. *Current Computer Aided-Drug Design*, 7(2), 146–157. <https://doi.org/10.2174/157340911795677602>
 28. Murai, S., Matuszkiewicz, J., Okuzono, Y., Miya, H., & de Jong, R. (2017). Aurora B inhibitor TAK-901 synergizes with bcl-xl inhibition by inducing active BAX in cancer cells. *Anticancer Research*, 37(2), 437–444. <https://doi.org/10.21873/anticancer.11335>
 29. Petros, A. M., Medek, A., Nettesheim, D. G., Kim, D. H., Yoon, H. S., Swift, K., Matayoshi, E. D., Oltersdorf, T., & Fesik, S. W. (2001). Solution structure of the antiapoptotic protein bcl-2. *Proceedings of the National Academy of Sciences*, 98(6), 3012–3017. <https://doi.org/10.1073/pnas.041619798>
 30. Qian, S., Wang, Z., Yang, W., Huang, J., Yang, Y., & Wang, J. (2022). The role of bcl-2 family proteins in regulating apoptosis and cancer therapy. *Frontiers in Oncology*, 12. <https://doi.org/10.3389/fonc.2022.985363>
 31. Roset, Ramon & Ortet, Laura & Gil-Gómez, Gabriel. (2007). Role of Bcl-2 family members on apoptosis: what we have learned from knock-out mice. *Frontiers in bioscience : a journal and virtual library*. 12. 4722-30.

32. Sattler, M., Liang, H., Nettlesheim, D., Meadows, R. P., Harlan, J. E., Eberstadt, M., Yoon, H., Shuker, S. B., Chang, B. S., Minn, A. J., Thompson, C. B., & Fesik, S. W. (1997). *STRUCTURE OF BCL-XL/BAK PEPTIDE COMPLEX, NMR, MINIMIZED AVERAGE STRUCTURE*. <https://doi.org/10.2210/pdb1bxl/pdb>
33. Souers, A., Levenson, J., Boghaert, E., Ackler, S., Catron, N., Chen, J., ... & Elmore, S. (2013). Abt-199, a potent and selective bcl-2 inhibitor, achieves antitumor activity while sparing platelets. *Nature Medicine*, 19(2), 202-208. <https://doi.org/10.1038/nm.3048>
34. Studer, G., Biasini, M., & Schwede, T. (2014). Assessing the local structural quality of transmembrane protein models using statistical potentials (QMEANBrane). *Bioinformatics*, 30(17), i505–i511. <https://doi.org/10.1093/bioinformatics/btu457>
35. Thomadaki, H., Talieri, M., & Scorilas, A. (2006). Treatment of MCF-7 cells with taxol and etoposide induces distinct alterations in the expression of apoptosis-related genes BCL2, BCL2L12, BAX, CASPASE-9 and FAS. *Biological Chemistry*, 387(8). <https://doi.org/10.1515/bc.2006.133>
36. Trott, O., & Olson, A. J. (2009). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31(2), 455–461. <https://doi.org/10.1002/jcc.21334>
37. Wani, M. C., Taylor, H. L., Wall, M. E., Coggon, P., & McPhail, A. T. (1971). Plant antitumor agents. VI. Isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. *Journal of the American Chemical Society*, 93(9), 2325–2327. <https://doi.org/10.1021/ja00738a045>
38. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, 46(W1), W296–W303. <https://doi.org/10.1093/nar/gky427>
39. Welch, C., Santra, M. K., El-Assaad, W., Zhu, X., Huber, W. E., Keys, R. A., Teodoro, J. G., & Green, M. R. (2009). Identification of a protein, G0S2, that lacks bcl-2 homology domains and interacts with and antagonizes bcl-2. *Cancer Research*, 69(17), 6782–6789. <https://doi.org/10.1158/0008-5472.can-09-0128>
40. Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., Assempour, N., Iynkkaran, I., Liu, Y., Maciejewski, A., Gale, N., Wilson, A., Chin, L., Cummings, R., Le, D., ... Wilson, M. (2017). DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Research*, 46(D1), D1074–D1082. <https://doi.org/10.1093/nar/gkx1037>
41. Xu, J., Jiao, F., & Berger, B. (2005). A tree-decomposition approach to protein structure prediction. *2005 IEEE Computational Systems Bioinformatics Conference (CSB'05)*. <http://dx.doi.org/10.1109/csb.2005.9>
42. Yang, C., Chen, E. A., & Zhang, Y. (2022). Protein–Ligand docking in the machine-learning

era. *Molecules*, 27(14), 4568. <https://doi.org/10.3390/molecules27144568>

43. Yap, J., Chen, L., Lanning, M., & Fletcher, S. (2016). Expanding the cancer arsenal with targeted therapies: disarmament of the antiapoptotic bcl-2 proteins by small molecules. *Journal of Medicinal Chemistry*, 60(3), 821-838. <https://doi.org/10.1021/acs.jmedchem.5b01888>